

**BLAST Basic Local Alignment Search Tool**SEQ 2

Job Title: BD205212:Nucleotide sequence for detecting...

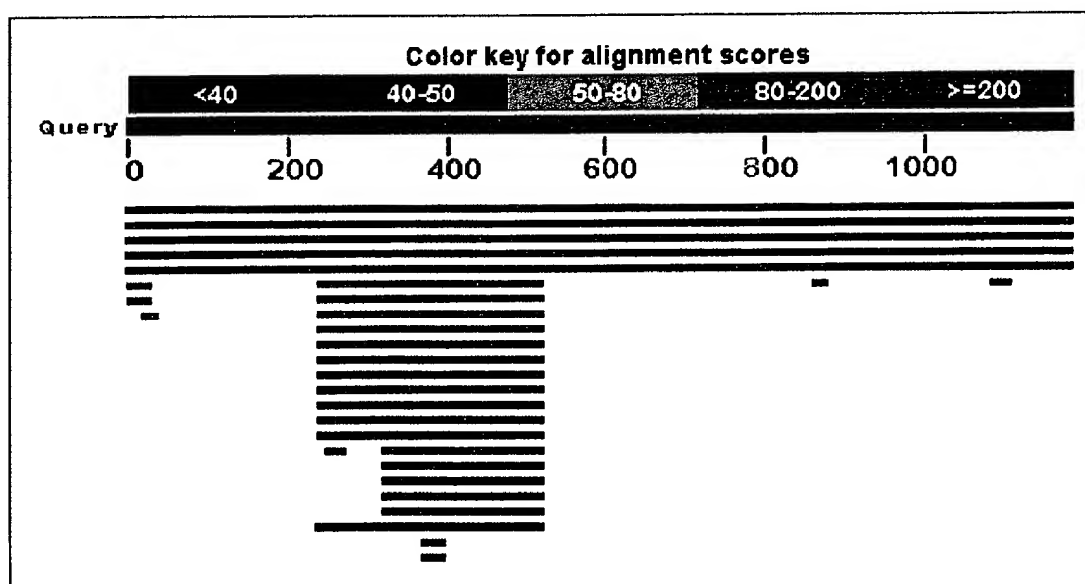
Document 2

Please, try our new design!

**BLASTN 2.2.18+**

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402. RID: 88NSGD54014 Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS, environmental samples or phase 0, 1 or 2 HTGS sequences) 7,064,549 sequences; 24,305,219,031 total letters

Query= gi|33014982|dbj|BD205212.1| Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC) Length=1181

**Distribution of 30 Blast Hits on the Query Sequence**

Distance tree of results **NEW**Legend for links to other resources: **U** UniGene **E** GEO **G** Gene **S** Structure**Sequences producing significant alignments:**

(Click headers to sort columns)

<b>AF401292.1</b>	Escherichia coli O157:H- plasmid pSF0157, complete sequence	2111	2111	100%	0.0	99%	
<b>AF074613.1</b>	Escherichia coli O157:H7 plasmid p0157, complete sequence	2111	2111	100%	0.0	99%	
<b>Y11275.1</b>	E.coli 7.4 kb DNA from plasmid p0157	2111	2111	100%	0.0	99%	<b>G</b>
<b>AB011549.2</b>	Escherichia coli O157:H7 str. Sakai plasmid p0157 DNA, complete sequence	2111	2111	100%	0.0	99%	
<b>AF043470.1</b>	Escherichia coli plasmid p0157 ecf4 gene, partial cds; and ecf3, ecf2, and ecf1 genes, complete cds	2102	2102	100%	0.0	99%	<b>G</b>
<b>CP000035.1</b>	Shigella dysenteriae Sd197 plasmid pSD1_197, complete sequence	138	138	24%	6e-29	71%	
<b>CP001064.1</b>	Escherichia coli 53638 plasmid p53638_226, complete sequence	132	132	24%	3e-27	71%	
<b>CP001062.1</b>	Shigella boydii CDC 3083-94 plasmid pBS512_211, complete sequence	132	132	24%	3e-27	71%	
<b>AF386526.1</b>	Shigella flexneri 2a str. 301 virulence plasmid pCP301, complete sequence	132	132	24%	3e-27	71%	
<b>AY206446.1</b>	Shigella flexneri plasmid pINV_F6 M1382 ORF186 (ORF186), VirK (virK), and MsbB2 (msbB2) genes, complete cds	132	132	24%	3e-27	71%	
<b>AL391753.1</b>	Shigella flexneri virulence plasmid pWR100: from 1 to 213494	132	132	24%	3e-27	71%	
<b>AF348706.1</b>	Shigella flexneri 5a plasmid virulence plasmid pWR501, complete sequence	132	132	24%	3e-27	71%	
<b>AY879342.1</b>	Shigella flexneri plasmid pSF5, complete sequence	132	132	24%	3e-27	71%	
<b>CP000037.1</b>	Shigella boydii Sb227 plasmid pSB4_227, complete sequence	132	132	24%	3e-27	71%	
<b>CP000039.1</b>	Shigella sonnei Ss046 plasmid pSS_046, complete sequence	132	132	24%	3e-27	71%	
<b>D11025.1</b>	Shigella flexneri plasmid pMYSH6000 virK gene for virulence protein, complete cds	132	132	24%	3e-27	71%	
<b>CP000799.1</b>	Escherichia coli E24377A plasmid pETEC_74, complete sequence	111	111	17%	9e-21	71%	
<b>CP000795.1</b>	Escherichia coli E24377A plasmid pETEC_80, complete sequence	111	111	17%	9e-21	71%	
<b>AB255435.1</b>	Escherichia coli plasmid p086A1 DNA, complete sequence	111	111	17%	9e-21	71%	
<b>AF134403.1</b>	Escherichia coli plasmid pAA2 Shf (shf), hexosyltransferase homolog (capU), and VirK (virK) genes, complete cds	111	111	17%	9e-21	71%	
<b>CR942285.1</b>	Escherichia coli plasmid pCoo	111	111	17%	9e-21	71%	
<b>CP000800.1</b>	Escherichia coli E24377A, complete genome	107	107	24%	1e-19	69%	
<b>AC133908.7</b>	Mus musculus chromosome 5, clone RP24-299L9, complete sequence	46.4	46.4	2%	0.30	93%	
<b>AC127327.4</b>	Mus musculus BAC clone RP23-254M18 from	46.4	46.4	2%	0.30	93%	

## 5, complete sequence

<b>XU025714.1</b>	<i>Salmo salar</i> retinoic acid receptor gamma a (Rarga), coiled-coil transcriptional coactivator a (Kiaa1536), homeobox protein HoxC13aa (HoxC13aa), homeobox protein HoxC12aa (HoxC12aa), homeobox protein HoxC11aa (HoxC11aa), homeobox protein HoxC10aa (HoxC10aa), homeobox protein HoxC9aa (HoxC9aa), homeobox protein HoxC8aa (HoxC8aa), and homeobox protein HoxC6aa (HoxC6aa) genes, complete cds	44.6	44.6	2%	1.1	93%	
<b>BX571861.1</b>	<i>Photorhabdus luminescens</i> subsp. laumondii TT01 complete genome; segment 3/17	44.6	44.6	2%	1.1	100%	
<b>DQ192243.1</b>	<i>Operophtera brumata</i> reovirus segment 9, complete sequence	42.8	42.8	2%	3.7	92%	<b>G</b>
<b>AC145866.3</b>	Pan troglodytes BAC clone RP43-21B7 from chromosome 7, complete sequence	42.8	42.8	2%	3.7	90%	
<b>AE015928.1</b>	<i>Bacteroides thetaiotaomicron</i> VPI-5482, complete genome	42.8	42.8	1%	3.7	100%	
<b>AC004844.1</b>	<i>Homo sapiens</i> PAC clone RP4-613I23 from 7p11-p13, complete sequence	42.8	42.8	2%	3.7	90%	<b>E</b>

**Alignments**

>gb|AF401292.1| **D** Escherichia coli O157:H- plasmid pSF0157, complete sequence  
Length=121239

Score = 2111 bits (2340), Expect = 0.0  
Identities = 1178/1182 (99%), Gaps = 1/1182 (0%)  
Strand=Plus/Minus

Query	1	CTGCAGGAGA-TGG	TTGCCCATCCCAGCGCGCTCCAGCT	59
Sbjct	29247	CTGCAGGAGAGTGGAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCT		29
Query	60	GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCGTATTTAAACAAT		11
Sbjct	29187	GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCGTATTTAAACAAT		29
Query	120	GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTTCGGGCTGAAAAGAGGATC		17
Sbjct	29127	GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTTCGGGCTGAAAAGAGGATC		29
Query	180	CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC		23
Sbjct	29067	CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC		29
Query	240	GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG		29
Sbjct	29007	GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG		28
Query	300	AAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCC		35
Sbjct	28947	AAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCC		28
Query	360	ACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA		41
Sbjct	28887	ACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA		28
Query	420	CTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCA		47
Sbjct	28827	CTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCA		28
Query	480	TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGG		53
Sbjct	28767	TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGG		28
Query	540	GCACCTGCCATTAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGT		59
Sbjct	28707	GCACCTGCCATTAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGT		28
Query	600	GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCCTTCCATC		65
Sbjct	28647	GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCCTTCCATC		28
Query	660	TTCAGCTGATACAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCC		71
Sbjct	28587	TTCAGCTGATACAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCC		28
Query	720	ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGT		77
Sbjct	28527	ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGT		28
Query	780	AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAATAACCGTT		83
Sbjct	28467	AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAATAACCGTT		28
Query	840	ACCGGTGTCCTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA		89
Sbjct	28407	ACCGGTGTCCTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA		28
Query	900	ATGAAATCACACAGATAATTACGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT		95
Sbjct	28347	ATGAAATCACACAGATAATTACGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT		28
Query	960	TGTCTGACAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTG		10

```

Sbjct  28287  |||||
TGTCGTGACAATACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTG 28
Query  1020    TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTCCTCGATAATCCGTCGAAGT 10
Sbjct  28227    TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTCCTCGATAATCCGTCGAAGT 28
Query  1080    CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGGAATAAAAGTGACATCATGCCCT 11
Sbjct  28167    CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGGAATAAAAGTGACATCATGCCCT 28
Query  1140    CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
Sbjct  28107    CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG 28066

```

>gb|AF074613.1| **D** Escherichia coli O157:H7 plasmid pO157, complete sequence  
Length=92077

Score = 2111 bits (2340), Expect = 0.0  
Identities = 1178/1182 (99%), Gaps = 1/1182 (0%)  
Strand=Plus/Minus

```

Query  1      CTGCAGGAGA-TGGGAAAAAGGCTGAAAGGAAAAATTGCCCATCCCAGCGCGCTCCAGCT 59
Sbjct  19435     CTGCAGGAGAGTGAAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCT 19
Query  60      GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAAT 11
Sbjct  19375     GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAAT 19
Query  120     GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTCTGGGCTGAAAAGAGGATC 17
Sbjct  19315     GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTCTGGGCTGAAAAGAGGATC 19
Query  180     CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC 23
Sbjct  19255     CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC 19
Query  240     GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG 29
Sbjct  19195     GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG 19
Query  300     AAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCC 35
Sbjct  19135     AAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCC 19
Query  360     ACGGGTAACAGCGTCCCTGTACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA 41
Sbjct  19075     ACGGGTAACAGCGTCCCTGTACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA 19
Query  420     CTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCA 47
Sbjct  19015     CTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCA 18
Query  480     TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGG 53
Sbjct  18955     TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGG 18
Query  540     GCACCTGCCATTAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGT 59
Sbjct  18895     GCACCTGCCATTAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGT 18
Query  600     GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCTTCCATC 65
Sbjct  18835     GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCTTCCATC 18
Query  660     TTCAGCTGATACAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCC 71
Sbjct  18775     TTCAGCTGATACAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCC 18
Query  720     ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCAGGCGTGGATATGCAGGGGT 77
Sbjct  18715     ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCAGGCGTGGATATGCAGGGGT 18
Query  780     AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTT 83
Sbjct  18655     AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTT 18
Query  840     ACCGGTGTCTTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA 89

```

01/07/2020

```

Sbjct  2166  |||||
AACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCAGGCGTGGATATGCAGGGGT 210
Query  780  AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTT 839
Sbjct  2106  |||||
AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTT 204
Query  840  ACCGGTGTCTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA 899
Sbjct  2046  |||||
ACCGGTGTCTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGCTAGGCACAACA 198
Query  900  ATGAAATCACACAGATAATTTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT 959
Sbjct  1986  |||||
ATGAAATCACACAGATAATTTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT 192
Query  960  TGTCTGACAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTG 101
Sbjct  1926  |||||
TGTCTGACAATACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTG 186
Query  1020  TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGT 107
Sbjct  1866  |||||
TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGT 180
Query  1080  CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGGAATAAAAGTGACATCATGCCCT 113
Sbjct  1806  |||||
CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGGAATAAAAGTGACATCATGCCCT 174
Query  1140  CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
Sbjct  1746  |||||
CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1705

```

>dbj|AB011549.2| **D** Escherichia coli O157:H7 str. Sakai plasmid pO157 DNA, complete sequence  
Length=92721

Score = 2111 bits (2340), Expect = 0.0  
Identities = 1178/1182 (99%), Gaps = 1/1182 (0%)  
Strand=Plus/Minus

```

Query  1  CTGCAGGAGA-TGG-CTGCAGGAGAGTGGATTGCCCATCCCAGCGCGCTCCAGCT 59
Sbjct  88950  CTGCAGGAGAGTGGAAAAAAAGCCAAAATAAAAAATTGCCCATCCCAGCGCGCTCCAGCT 88
Query  60  GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAAT 11
Sbjct  88890  GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCCGTATTTAAACAAT 88
Query  120  GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTCCGGGTGAAAAGAGGATC 17
Sbjct  88830  GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTCCGGGTGAAAAGAGGATC 88
Query  180  CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC 23
Sbjct  88770  CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC 88
Query  240  GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG 29
Sbjct  88710  GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG 88
Query  300  AAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCC 35
Sbjct  88650  AAGCGTTCGGTTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCC 88
Query  360  ACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA 41
Sbjct  88590  ACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA 88
Query  420  CTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCA 47
Sbjct  88530  CTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCA 88
Query  480  TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGG 53
Sbjct  88470  TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGGTAACGCTGGGAAAAGG 88
Query  540  GCACCTGCCATTAACACATCTCCGCTCATTTCCAGGTGTTCTGTCTGCTGACGCAGACGT 59
Sbjct  88410  GCACCTGCCATTAACACATCTCCGCTCATTTCCAGGTGTTCTGTCTGCTGACGCAGACGT 88

```

```

Query 600      GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCCTTCCATC 65
                |||
Sbjct 88350    GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCCTTCCATC 88

Query 660      TTCAGCTGATACAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCC 71
                |||
Sbjct 88290    TTCAGCTGATACAATACACGCAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCC 88

Query 720      ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGT 77
                |||
Sbjct 88230    ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGT 88

Query 780      AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAAATAACCGTT 83
                |||
Sbjct 88170    AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAAATAACCGTT 88

Query 840      ACCGGTGTCTTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA 89
                |||
Sbjct 88110    ACCGGTGTCTTGACACCTTCCGCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA 88

Query 900      ATGAAATCACACAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT 95
                |||
Sbjct 88050    ATGAAATCACACAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT 87

Query 960      TGTCTGACAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCTGCTATATTACTG 10
                |||
Sbjct 87990    TGTCTGACAATACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTCTGCTATATTACTG 87

Query 1020     TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATCCCCGATAATCCGTCGAAGT 10
                |||
Sbjct 87930    TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATCCCCGATAATCCGTCGAAGT 87

Query 1080     CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTGAAAGGAATAAAAGTGACATCATGCCCT 11
                |||
Sbjct 87870    CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTGAAAGGAATAAAAGTGACATCATGCCCT 87

Query 1140     CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
                |||
Sbjct 87810    CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG 87769

```

>gb|AF043470.1|AF043470 **G** Escherichia coli plasmid pO157 ecf4 gene, partial c  
ecf2, and ecf1 genes, complete cds  
Length=5612

GENE ID: 5290933 SF0157 p16 | w0016 [Escherichia coli]  
(10 or fewer PubMed links)

Score = 2102 bits (2330), Expect = 0.0  
Identities = 1176/1182 (99%), Gaps = 1/1182 (0%)  
Strand=Plus/Plus

```

Query 1        CTGCAGGAGA-TGGTTGCCCATCCCAGCGCGCTCCAGCT 59
                |||
Sbjct 2404     CTGCAGGAGAGTGGAAGGAGGCAAAATGAAAATTTGCCCATCCCAGCGCGCTCCAGCT 246

Query 60       GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCGTATTTAAACAAT 119
                |||
Sbjct 2464     GAAAGTAGGCCTGTTCTGTCCGGTATTTAAATGCATTGACCGTCCCGTATTTAAACAAT 252

Query 120      GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATC 179
                |||
Sbjct 2524     GTGATAAATTACTCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATC 258

Query 180      CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC 239
                |||
Sbjct 2584     CGCCGTTATCTGTTGCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCC 264

Query 240      GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG 299
                |||
Sbjct 2644     GTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGG 270

Query 300      AAGCGTTCCGGTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCC 359
                |||
Sbjct 2704     AAGCGTTCCGGTCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCC 276

Query 360      ACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA 419
                |||
Sbjct 2764     ACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCA 282

```



```

Query  420      CTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACCAAACGCTTCA  479
          |||
Sbjct  2824      CTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACCAAACGCTTCA  288

Query  480      TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCCGGTAACGCTGGGAAAAGG  539
          |||
Sbjct  2884      TTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCCGGTAACGCTGGGAAAAGG  294

Query  540      GCACCTGCCATTAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACGT  599
          |||
Sbjct  2944      GCACCTGCCATTAACACATCTCCGCTCATTCCCAGGTGTTCTGTCTGCTGACGCAGACTT  300

Query  600      GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCCTTCCATC  659
          |||
Sbjct  3004      GCTTCGTATTCTTCACGCCCGGCGCCACCACGAGCCAGCGAAATGATTTCCCTTCCATC  306

Query  660      TTCAGCTGATACAATACACGCAGCATAAATTTCATGTCCTTTTTCGGGACGTAGCATCCCC  719
          |||
Sbjct  3064      TTCAGCTGATACAATACACGCAGCATAAATTTCATGTCCTTTTTCGGGACGTAGCATCCCC  312

Query  720      ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGT  779
          |||
Sbjct  3124      ACCTGAACGATAAGCGGAACATTGTCTGCTGATGCAGCCCAGGCGTGGATATGCAGGGGT  318

Query  780      AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTT  839
          |||
Sbjct  3184      AACGGTCGCATGGCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGAATAACCGTT  324

Query  840      ACCGGTGTCTCTGACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGATAGGCACAACA  899
          |||
Sbjct  3244      ACCGGTGTCTCTGACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGCTAGGCACAACA  330

Query  900      ATGAAATCACACAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT  959
          |||
Sbjct  3304      ATGAAATCACACAGATAATTCAGGGAAAACGTTCTGGTCTTACGGGTGATGTAGGTTTTT  336

Query  960      TGTCTGACAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTG  101
          |||
Sbjct  3364      TGTCTGACAATACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTG  342

Query  1020     TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGT  107
          |||
Sbjct  3424      TCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTCCCCGATAATCCGTCGAAGT  348

Query  1080     CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGGAATAAAAGTGACATCATGCCCT  113
          |||
Sbjct  3484      CTGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGGAATAAAAGTGACATCATGCCCT  354

Query  1140     CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG  1181
          |||
Sbjct  3544      CTTTTTCTGGCTTCCGGAGCAATTTTACTTTTTTCTCTGCAG  3585

```

>gb|CP000035.1| **D** Shigella dysenteriae Sd197 plasmid pSD1\_197, complete sequence  
Length=182726

Score = 138 bits (152), Expect = 6e-29  
Identities = 207/289 (71%), Gaps = 4/289 (1%)  
Strand=Plus/Minus

```

Query  237      GCCGTTCTGTAAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCAT  29
          |||
Sbjct  87991      GCAGTTTTTATTAATATCAAACCGCTCTTCTATATCCTGTTTTGC-CTGGCGAGCCATCTG  87

Query  296      CCGGAAGCGTTCGG-TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACAT  35
          |||
Sbjct  87932      ATAAAAACGTCCAGGGTCATTA-AAAAATCATTACGGGCGCACATCCATGCGTGCTTAT  87

Query  355      CCCCCACGGGTAACAGCGTCCCTGTACATTCTTCTGAATGACATCAGGGATCCCGCCCG  41
          |||
Sbjct  87873      TACCTGCTGGTAACAATGTCCCGGTCTGGTTGTTCTGAATAACATCAGGGATTCACCAA  87

Query  415      TCTCACTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACCAAACG  47
          |||
Sbjct  87813      TCTGACTGGCCAGTACAGGCACAGAAAATGCCGATGCTTCTGCCAGTACCATAACCAAACG  87

Query  475      CTTCAATTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGG  523
          |||
Sbjct  87753      ATTCGTTTTTCTGAAGGCAGAACACCAGACTGGCAACCCGATATACGGG  87705

```

>gb|CP001064.1| **D** Escherichia coli 53638 plasmid p53638\_226, complete sequence  
Length=225683

Features in this part of subject sequence:  
**UDP-sugar hydrolase**

Score = 132 bits (146), Expect = 3e-27  
Identities = 206/289 (71%), Gaps = 4/289 (1%)  
Strand=Plus/Plus

```
Query 237      GCCGTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCAT 2
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 115503    GCAGTTTATTATTAATATCAAACCGCTCTTCTATATCCTGTTTTCG-CTGGCGAGCCATCTG 1

Query 296      CCGGAAGCGTTCCGG-TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACAT 3
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 115562    ATAAAAACGCCAGGGTCATTA-AAAAAATCATTCAGGGCGCACATCCATGCGTGCTTAT 1

Query 355      CCCCCACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCG 4
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 115621    TACCTGCTGGTAACAATGTCCCGGTCTGGTTGTTCTGAATAACATCAGGGATTCCACCAA 1

Query 415      TCTCACTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACAAACG 4
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 115681    TCTGACTGGCCAGTACAGGCACAGAAAATGCCGATGCTTCTGCCAGTACCATAACAAACG 1

Query 475      CTCATTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGG 523
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 115741    ATTCGTTTTCTGAAGGCAGAACCACCAGACTGGCAACCCGATATACGGG 115789
```

>gb|CP001062.1| **D** Shigella boydii CDC 3083-94 plasmid pBS512\_211, complete sequence  
Length=210919

Features in this part of subject sequence:  
**putative glycosyl transferase, group 1 family protein**

Score = 132 bits (146), Expect = 3e-27  
Identities = 206/289 (71%), Gaps = 4/289 (1%)  
Strand=Plus/Minus

```
Query 237      GCCGTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCAT 2
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 176254    GCAGTTTATTATTAATATCAAACCGCTCTTCTATATCCTGTTTTCG-CTGGCGAGCCATCTG 1

Query 296      CCGGAAGCGTTCCGG-TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACAT 3
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 176195    ATAAAAACGCCAGGGTCATTA-AAAAAATCATTCAGGGCGCACATCCATGCGTGCTTAT 1

Query 355      CCCCCACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCG 4
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 176136    TACCTGCTGGTAACAATGTCCCGGTCTGGTTGTTCTGAATAACATCAGGGATTCCACCAA 1

Query 415      TCTCACTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATAACAAACG 4
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 176076    TCTGACTGGCCAGTACAGGCACAGAAAATGCCGATGCTTCTGCCAGTACCATAACAAACG 1

Query 475      CTCATTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGG 523
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 176016    ATTCGTTTTCTGAAGGCAGAACCACCAGACTGGCAACCCGATATACGGG 175968
```

>gb|AF386526.1| **D** Shigella flexneri 2a str. 301 virulence plasmid pCP301, complete sequence  
Length=221618

Features in this part of subject sequence:  
**UDP-sugar hydrolase**

Score = 132 bits (146), Expect = 3e-27  
Identities = 206/289 (71%), Gaps = 4/289 (1%)  
Strand=Plus/Minus


```
Query 237      GCCGTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCAT 2
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 193608    GCAGTTTATTATTAATATCAAACCGCTCTTCTATATCCTGTTTTCG-CTGGCGAGCCATCTG 1

Query 296      CCGGAAGCGTTCCGG-TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACAT 3
```

21/07/2008



Sbjct 97908 ATTCGTTTCTGAAGGCAGAACACCAGACTGGCAACCCGATATACGGG 97860

```
>gb|CP000039.1|  Shigella sonnei Ss046 plasmid pSS_046, complete sequence
Length=214396
```

Features in this part of subject sequence:  
conserved hypothetical protein

Score = 132 bits (146), Expect = 3e-27  
Identities = 206/289 (71%), Gaps = 4/289 (1%)  
Strand=Plus/Minus

Query	237	GCCGTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCAT	2
Sbjct	158570	GCAGTTTTATTAATATCAAACCGCTCTTCTATATCCTGTTTTCG-CTGGCGAGCCATCTG	1
Query	296	CCGGAAGCGTTCGGG-TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACAT	3
Sbjct	158511	ATAAAAACGCCCAGGGTCATTA-AAAAAATCATTAGGGCGCACATCCATGCGTGCTTAT	1
Query	355	CCCCACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCCG	4
Sbjct	158452	TACCTGCTGGTAACAATGTCCCGGTCTGGTTGTTCTGAATAACATCAGGGATTCCACCAA	1
Query	415	TCTCACTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATACCAAACG	4
Sbjct	158392	TCTGACTGGCCAGTACAGGCACAGAAAATGCCGATGCTTCTGCCAGTACCATACCAAACG	1
Query	475	CTTCATTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGG	523
Sbjct	158332	ATTCGTTTTCTGAAGGCAGAACACCAGACTGGCAACCCGATATACGGG	158284

```
>dbj|D11025.1|SHFVIRK Shigella flexneri plasmid pMYSH6000 virK gene for virul.
complete cds
Length=1642
```

Score = 132 bits (146), Expect = 3e-27  
Identities = 206/289 (71%), Gaps = 4/289 (1%)  
Strand=Plus/Minus


Query	237	GCCGTTCTGTTAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCAT	295
Sbjct	334	GCAGTTTTATTAAATATCAAACCGCTCTTCTATATCCTGTTTTGC-CTGGCGAGCCATCTG	276
Query	296	CCGGAAGCGTTCGGG-TCGGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACAT	354
Sbjct	275	ATAAAAACGCCCAGGGTCATTA-AAAAAATCATTCAGGGCGCACATCCATGCGTGCTTAT	217
Query	355	CCCCACGGGTAACAGCGTCCCTGTCTCACATTCTTCTGAATGACATCAGGGATCCC GCCCG	414
Sbjct	216	TACCTGCTGGTAACAATGTCCCGTCTGTTATTCTGAATAACATCAGGGATTCCACCAA	157
Query	415	TCTCACTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATAACCAAACG	474
Sbjct	156	TCTGACTGGCCAGTACAGGCACAGAAAAATGCCGATGCTTCTGCCAGTACCATAACCAAACG	97
Query	475	CTTCATTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGG	523
Sbjct	96	ATTGCTTTTCTGAAGGCAGAACCACCAGACTGGCAACCCGATATACGGG	48

```
>gb|CP000799.1|  Escherichia coli E24377A plasmid pETEC_74, complete sequence  
Length=74224
```

Score = 111 bits (122), Expect = 9e-21  
Identities = 148/206 (71%), Gaps = 0/206 (0%)  
Strand=Plus/Minus


Query	318	AAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCT	37
Sbjct	61325	AAAAAATCATTACAGGGCGCACATCCATGCGTGCTTATTACCTGCTGGTAACAATGTCCCG	61
Query	378	GTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAACGGGCACG	43
Sbjct	61265	GTCTGGTTGTTCTGAATAACCTCAGGGATTCCACCAATCTGAGTGGCCACTACAGGCACA	61
Query	438	CCGGAGACTGACGCTTCAGCCAGTACCATAACCAACGCTTCATTTCGAAGGCATGACC	49

Sbjct	61205	GAAATGCGGATGCTTCTGCCAGCACCATACCAAAAGATTTCGTTTTCTGAAGGCAGAACC	61
Query	498	ACCACACTGGCAATCCGGTAGACCGG	523
Sbjct	61145	ACCAGACTGGCAACCCGATATACGGG	61120

```
>gb|CP000795.1|  Escherichia coli E24377A plasmid pETEC_80, complete sequence  
Length=79237
```

Score = 111 bits (122), Expect = 9e-21  
Identities = 148/206 (71%), Gaps = 0/206 (0%)  
Strand=Plus/Plus

Query	318	AAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCACGGGTAACAGCGTCCCT	37
Sbjct	37177	AAAAAATCATTTCAGGGCGCACATCCATGCGTGCTTATTACCTGCTGGTAACAATGTCCCG	37
Query	378	GTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAACGGGCACG	43
Sbjct	37237	GTCTGGTTGTTCTGAATAACCTCAGGGATTCCACCAATCTGAGTGGCCACTACAGGCACA	37
Query	438	CCGGAGACTGACGCTTCAGCCAGTACCATAACCAAACGCTTCATTTTCCGAAGGCATGACC	49
Sbjct	37297	GAAAATGCCGATGCTTCTGCCAGCACCATACCAAAAAGATTTCGTTTTCTGAAGGCAGAACC	37
Query	498	ACCACACTGGCAATCCGGTAGACCGG	523
Sbjct	37357	ACCAGACTGGCAACCCGATATACGGG	37382

```
>dbj|AB255435.1|  Escherichia coli plasmid pO86A1 DNA, complete sequence
Length=120730
```

Score = 111 bits (122), Expect = 9e-21  
Identities = 148/206 (71%), Gaps = 0/206 (0%)  
Strand=Plus/Plus

Query	318	AAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCT	377
Sbjct	2154	AAAAAATCATTTCAGGGCGCACATCCATGCGTGCTTATTACCTGCTGGTAACAATGTCCCG	221
Query	378	GTCACATTCTTCTGAATGACATCAGGGATCCGCCCGTCTCACTGGCGATAACGGGCACG	437
Sbjct	2214	GTCTGGTTGTTCTGAATAACCTCAGGGATTCCACCAATCTGAGTGGCCACTACAGGCACA	227
Query	438	CCGGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCATTTTCCGAAGGCATGACC	497
Sbjct	2274	GAAAATGCCGATGCTTCTGCCAGCACCATACCAAAAGATTCTGTTTTCTGAAGGCAGAACC	233
Query	498	ACCACACTGGCAATCCGGTAGACCGG	523
Sbjct	2334	ACCAGACTGGCAACCCGATATACGGG	2359

>**gb|AF134403.1|AF134403** Escherichia coli plasmid pAA2 Shf (shf), hexosyltrans (capU), and VirK (virK) genes, complete cds  
Length=3500

Score = 111 bits (122), Expect = 9e-21  
Identities = 148/206 (71%), Gaps = 0/206 (0%)  
Strand=Plus/Minus

Query	318	AAAAAATCGCGCAGTGC	CGCGGTCCATGCAGACACATCCCCACGGGTAACAGCGTCCCT	377
Sbjct	2333	AAAAAATCATTCA	GGGCGCACATCCATGCGTGCTTATTACCTGCTGGTAACAATGTCCCG	227
Query	378	GTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAACGGGCACG	437	
Sbjct	2273	GTCTGGTTGTTCTGAATAACCTCAGGGATTCCACCAATCTGAGTGGCCACTACAGGCACA	221	
Query	438	CCGGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCATTTTCCGAAGGCATGACC	497	
Sbjct	2213	GAAAATGCCGATGCTTCTGCCAGCACCATAACCAAAAGATTTCGTTTTCTGAAGGCAGAACC	215	
Query	498	ACCACACTGGCAATCCGGTAGACCGG	523	
Sbjct	2153	ACCAGACTGGCAACCCGATATACGGG	2128	

>emb|CR942285.1| **D** Escherichia coli plasmid pCoo  
Length=98396

Score = 111 bits (122), Expect = 9e-21  
Identities = 148/206 (71%), Gaps = 0/206 (0%)  
Strand=Plus/Plus

```
Query 318      AAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACATCCCCACGGGTAACAGCGTCCCT 37
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 34153     AAAAAATCATTTCAGGGCGCACATCCATGCGTGCTTATTACCTGCTGGTAACAATGTCCCG 34
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 378      GTCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGATAACGGGCACG 43
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 34213     GTCTGGTTGTTCTGAATAACCTCAGGGATTCCACCAATCTGAGTGGCCACTACAGGCACA 34
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 438      CCGGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCATTTTCCGAAGGCATGACC 49
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 34273     GAAAATGCCGATGCTTCTGCCAGCACCATAACAAAAGATTGTTTTCTGAAGGCAGAACC 34
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 498      ACCACACTGGCAATCCGGTAGACCGG 523
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 34333     ACCAGACTGGCAACCCGATATACGGG 34358
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

>gb|CP000800.1| **D** Escherichia coli E24377A, complete genome  
Length=4979619

Features in this part of subject sequence:  
**putative glycosyl transferase, group 1 family protein**

Score = 107 bits (118), Expect = 1e-19  
Identities = 201/290 (69%), Gaps = 4/290 (1%)  
Strand=Plus/Plus

```
Query 236      TGCCGTTCTGT'TAATATCAAACCGGTACTCAATATCTTCTCTGGCGCTGGCT-GCCATCA
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 4845818   TGCAGTTTTAT'TAATATCAAACCGCTCTTCTATATCCTGTTTTGC-CAGGCGAGCCATCT
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 295      TCCGGAAGCGTTCCGG-TCCGGATAAAAAATCGCGCAGTGCGCCGGTCCATGCAGACACA
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 4845877   GATAAAAACGCCAGGGTCATTA-AAAAAATCATTAGGGCGCACATCCATGCGTGCTTA
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 354      TCCCCACGGGTAACAGCGTCCCTGTCACATTCTTCTGAATGACATCAGGGATCCCGCCC
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 4845936   TTACTTGCTGGTAACAATGTCCCGGTCTGGTTG'TTCTGAATAACCTCAGGGATTCCACCA
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 414      GTCTCACTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGCCAGTACCATAACAAAC
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 4845996   ATCTGAGTGGCCACTACAGGCACAGAAAATGCCGATGCTTCTGCCAGCACCATAACAAAA
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Query 474      GCTTCATTTTCCGAAGGCATGACCACCACACTGGCAATCCGGTAGACCGG 523
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 4846056   GATTCGTTTTCTGAAGGCAGAACCACCAGACTGGCAACCCGATATACGGG 4846105
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

>gb|AC133908.7| **D** Mus musculus chromosome 5, clone RP24-299L9, complete seque  
Length=214270

Score = 46.4 bits (50), Expect = 0.30  
Identities = 28/30 (93%), Gaps = 0/30 (0%)  
Strand=Plus/Minus

```
Query 4      CAGGAGATGGAAAAAAGCCAAAATAAAAA 33
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 185462   CAGGAGATGGAAAAAAGCTGAAATAAAAA 185433
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

>gb|AC127327.4| **D** Mus musculus BAC clone RP23-254M18 from 5, complete sequenc  
Length=198433

Score = 46.4 bits (50), Expect = 0.30  
Identities = 28/30 (93%), Gaps = 0/30 (0%)  
Strand=Plus/Minus

```
Query 4      CAGGAGATGGAAAAAAGCCAAAATAAAAA 33
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct 92692   CAGGAGATGGAAAAAAGCTGAAATAAAAA 92663
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

>gb|EU025714.1| **D** *Salmo salar* retinoic acid receptor gamma a (Rarga), coiled-transcriptional coactivator a (Kiaa1536), homeobox protein HoxC13aa (HoxC13aa), homeobox protein HoxC12aa (HoxC12aa), homeobox protein HoxC11aa (HoxC11aa), homeobox protein HoxC10aa (HoxC10aa), homeobox protein HoxC9aa (HoxC9aa), homeobox protein HoxC8aa (HoxC8aa), and homeobox protein HoxC6aa (HoxC6aa) genes, complete cds  
Length=212839

Features in this part of subject sequence:  
**retinoic acid receptor gamma a**

Score = 44.6 bits (48), Expect = 1.1  
Identities = 27/29 (93%), Gaps = 0/29 (0%)  
Strand=Plus/Plus

```
Query  1077  AGTCTGAGGATGGAAGGAAGGTGAAGGCT  1105
          ||| ||||| ||||| ||||| ||||| |||||
Sbjct  31660  AGTATGAGGATGGAAGGAAGGTGAGGGCT  31688
```

>emb|BX571861.1| **D** *Photorhabdus luminescens* subsp. *laumondii* T101 complete genome segment 3/17  
Length=342905

Features in this part of subject sequence:  
**unnamed protein product**

Score = 44.6 bits (48), Expect = 1.1  
Identities = 24/24 (100%), Gaps = 0/24 (0%)  
Strand=Plus/Minus

```
Query  20      agccaaaataaaaaattgccccatc  43
          ||||| ||||| ||||| ||||| |||||
Sbjct  27988  AGCCAAAATAAAAAATTGCCCATC  27965
```

>gb|DQ192243.1| **G** *Operophtera brumata* reovirus segment 9, complete sequence  
Length=1547

GENE ID: 5076725 OBRV\_s9gp1 | hypothetical protein  
[*Operophtera brumata* reovirus]

Score = 42.8 bits (46), Expect = 3.7  
Identities = 26/28 (92%), Gaps = 0/28 (0%)  
Strand=Plus/Minus

```
Query  247  TAATATCAAACCGGTACTCAATATCTTC  274
          ||||| ||||| ||||| ||||| |||||
Sbjct  260  TAATATCAAACCGGTATTCAATACCTTC  233
```

>gb|AC145866.3| **D** *Pan troglodytes* BAC clone RP43-21B7 from chromosome 7, complete sequence  
Length=189814

Score = 42.8 bits (46), Expect = 3.7  
Identities = 28/31 (90%), Gaps = 0/31 (0%)  
Strand=Plus/Plus

```
Query  366      AACAGCGTCCCTGTCACATTCTTCTGAATGA  396
          ||||| ||||| ||||| ||||| |||||
Sbjct  137128  AACAGAGTCCCTTTCACATCCTTCTGAATGA  137158
```

>gb|AE015928.1| **D** *Bacteroides thetaiotaomicron* VPI-5482, complete genome  
Length=6260361

Features in this part of subject sequence:  
**Histone-like bacterial DNA-binding protein**

Score = 42.8 bits (46), Expect = 3.7  
Identities = 23/23 (100%), Gaps = 0/23 (0%)  
Strand=Plus/Plus

```
Query  855      CCTTCGCCATCAGATGCGCCAT  877
```



Sbjct 6023752 ||||| CCTTCGCCATCAGATGCGCCAT 6023774

>gb|AC004844.1| **ED** Homo sapiens PAC clone RP4-613I23 from 7p11-p13, complete  
Length=141895

Score = 42.8 bits (46), Expect = 3.7  
Identities = 28/31 (90%), Gaps = 0/31 (0%)  
Strand=Plus/Plus

Query 366 AACAGCGTCCCTGTACAACTCTTCTGAATGA 396  
||||| ||||| ||||| ||||| |||||  
Sbjct 110633 AACAGAGTCCCTTTACATCCTTCTGAATGA 110663

Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS,  
GSS, environmental samples or phase 0, 1 or 2 HTGS sequences)

Posted date: Jul 20, 2008 5:48 PM

Number of letters in database: -1,464,584,741

Number of sequences in database: 7,064,549

Lambda K H  
0.634 0.408 0.912

Gapped

Lambda K H  
0.634 0.408 0.912

Matrix: blastn matrix:2 -3

Gap Penalties: Existence: 5, Extension: 2

Number of Sequences: 7064549

Number of Hits to DB: 12224137

Number of extensions: 604140

Number of successful extensions: 8350

Number of sequences better than 10: 0

Number of HSP's better than 10 without gapping: 0

Number of HSP's gapped: 8350

Number of HSP's successfully gapped: 0

Length of query: 1181

Length of database: 24305219031

Length adjustment: 36

Effective length of query: 1145

Effective length of database: 24050895267

Effective search space: 27538275080715

Effective search space used: 27538275080715

A: 0

X1: 22 (20.1 bits)

X2: 33 (29.8 bits)

X3: 110 (99.2 bits)

S1: 29 (27.4 bits)

S2: 45 (41.9 bits)